AI-Driven Advances in Early Pancreatic Cancer Detection

A Clinical and Technical Whitepaper on Radiomics, Deep Learning, and Risk Stratification

Prepared for: Al Herndon
Date: June 2026

> *Executive Summary: Pancreatic ductal adenocarcinoma (PDAC) remains one of the deadliest malignancies worldwide, largely due to its late-stage symptomatic clinical presentation. This report evaluates the paradigm shift toward AI-driven early interception, focusing on how advanced computer vision, radiomics feature engineering, and ensemble learning architectures isolate subvisual biomarkers up to 15 months before macroscopic lesions manifest on conventional scans.*

1. What: AI and Early Detection Paradigms

Pancreatic ductal adenocarcinoma (PDAC) is associated with an exceptionally poor prognosis, carrying an overall five-year survival rate of less than 10% globally. This clinical outcome is largely a function of the disease's insidious onset and a historical absence of specific early diagnostic symptoms. Consequently, over 85% of cases are diagnosed at an unresectable or widely metastatic stage, where therapeutic options are strictly limited to palliative care. However, localized tumors confined strictly to the primary tissue site demonstrate significantly better five-year relative survival rates approaching 40%, emphasizing that identifying early-stage resectable lesions represents the most viable path toward reducing national and global cancer mortality.

The Clinical Dilemma: Missed vs. Imaging-Occult Lesions

In current clinical practice, traditional multi-detector computed tomography (CT) and magnetic resonance imaging (MRI) frequently fail to identify pre-clinical pancreatic tumors. Medical literature distinguishes between two distinct diagnostic sub-categories: 'missed' pancreatic cancers and 'imaging-occult' pancreatic cancers. Missed lesions are those where a retrospective post-diagnostic expert review can visualize a discrete focal mass that was overlooked due to technical factors (such as suboptimal contrast timing or standard motion artifacts) or perceptual human errors. In stark contrast, imaging-occult cancers show no discernible macroscopic abnormalities or mass effects, even upon expert multidisciplinary re-review. In these instances, the presence of localized malignancy is encoded within subvisual parenchymal or architectural textural disruptions that escape standard human visual perception entirely.

Emerging AI Screening Modalities

To overcome the physical boundaries of human radiological review, current artificial intelligence frameworks are being deployed across three primary screening and diagnostic workflows:

- **Radiomics-Based Image Analysis:** Extracting high-throughput quantitative texture features from standard contrast-enhanced portal venous phase CT scans to uncover early indicators of carcinogenesis months before a mass forms.
- **Real-Time Computer-Aided Diagnosis (CAD):** Leveraging hybrid deep learning structures, such as unified Convolutional Neural Networks and Long Short-Term Memory (CNN-LSTM) networks, directly during Endoscopic Ultrasound (EUS) procedures to automatically characterize focal masses in real-time.
- **Electronic Health Record (EHR) Sequencing:** Using language processing concepts to sequence sequential health records (e.g., changes in international classification codes, weight loss indicators, and rapid blood glucose shifts) to identify a 25-fold risk spike up to 36 months prior to typical diagnosis.

2. How It Works: Technical Architectures and Pipelines

AI-driven diagnostic models operate through a series of automated, multi-stage pipelines designed to eliminate inter-observer manual variability and extract reproducible structural signals from medical data. This section breaks down the underlying engineering mechanics of these clinical tools.

The Radiomics and Image Processing Pipeline

Advanced frameworks, such as the Radiomics-based Early Detection MODel (REDMOD) and Project Felix, process standard abdominal CT scans using a standardized four-step computing pipeline:

Pipeline Stage	Technical Implementation & Engineering Mechanics
1. Automated Segmentation	Deep learning networks utilizing 3D nnU-Net architectures automatically isolate and trace the volumetric boundaries of the pancreas. This fully automated localization eliminates the labor-intensive errors and high variability associated with traditional manual contouring.
2. Preprocessing & Filtering	The segmented image volume undergoes standardized spatial resampling and gray-level normalization. Multi-scale digital image filters—specifically Wavelet transforms and Laplacian-of-Gaussian (LoG) filters—are applied to extract micro-architectural texture signals across multiple frequencies.
3. High-Throughput Feature Extraction	A vast panel of 968 quantitative radiomic features is extracted from the volumetric region of interest. These include first-order statistical metrics (e.g., voxel intensity, mean, entropy), shape-based indicators, and complex multi-directional gray-level matrices.
4. Feature Reduction & Selection	To resolve the high-dimensional 'curse of dimensionality' and eliminate non-informative noise, the Minimum Redundancy Maximum Relevance (mRMR) or Least Absolute Shrinkage and Selection Operator (LASSO) selection methods reduce the input to a core signature of the 40 most informative biomarkers.

Ensemble Classifiers and Addressing Class Imbalance

Because pancreatic cancer is a low-prevalence condition within the general population, training datasets exhibit an extreme class imbalance where negative control scans vastly outnumber pre-diagnostic cancer cases. To mitigate this majority-class bias and prevent catastrophic false-negative rates, algorithms employ the Synthetic Minority Over-sampling Technique (SMOTE). SMOTE mathematically generates synthetic pre-diagnostic instances within the feature space by interpolating between neighboring minority data points, creating a balanced 1:1 ratio for model optimization without simple data duplication.

The final classification engine utilizes a heterogeneous ensemble architecture that syndicates three distinct machine learning strategies: (a) Logistic Regression (LR) to establish well-calibrated, baseline linear probabilities; (b) Random Forest (RF) to capture complex, non-linear relationships via bootstrapped decision trees; and (c) Extreme Gradient Boosting (XGBoost) to minimize model residual errors sequentially. The outputs are integrated via a soft-voting mechanism that averages predicted probabilities to yield the final diagnostic outcome.

Real-Time CAD Modalities: Combining CNN and LSTM

For live diagnostic tracking during multi-parametric EUS examinations, a distinct hybrid neural network architecture is utilized. Standard deep learning models treat frame sequences as completely independent snapshots. To preserve critical diagnostic context, a unified architecture couples a Convolutional Neural Network (CNN) with a Long Short-Term Memory (LSTM) recurrent layer. The CNN operates as a spatial feature extractor, isolating discriminative properties across gray-scale, color Doppler, contrast-enhanced, and elastography phases. The LSTM layer then ingests these spatial matrices sequentially, successfully capturing temporal and state-dependent dependencies across the live ultrasound feed to distinguish chronic pancreatitis from highly malignant adenocarcinomas with an overall accuracy of 98.26%.

Clinical Screening and Risk Stratification Sieves

To maximize health system cost-effectiveness, AI deployment follows a multi-tiered diagnostic sieve (the DEF Approach). First, automated screening of massive electronic health record registries isolates clinical indicators—such as glycaemically-defined new-onset diabetes and rapid weight change evaluated through models like the ENDPAC score—to define a high-risk group. Second, serum or liquid biopsy biomarker testing further enriches this cohort. Third, the localized very high-risk group undergoes advanced automated radiological screening (e.g., REDMOD), enabling proactive clinical interception a median of 475 days prior to standard symptom-driven diagnosis.

Figure 1 What AI is doing in early detection

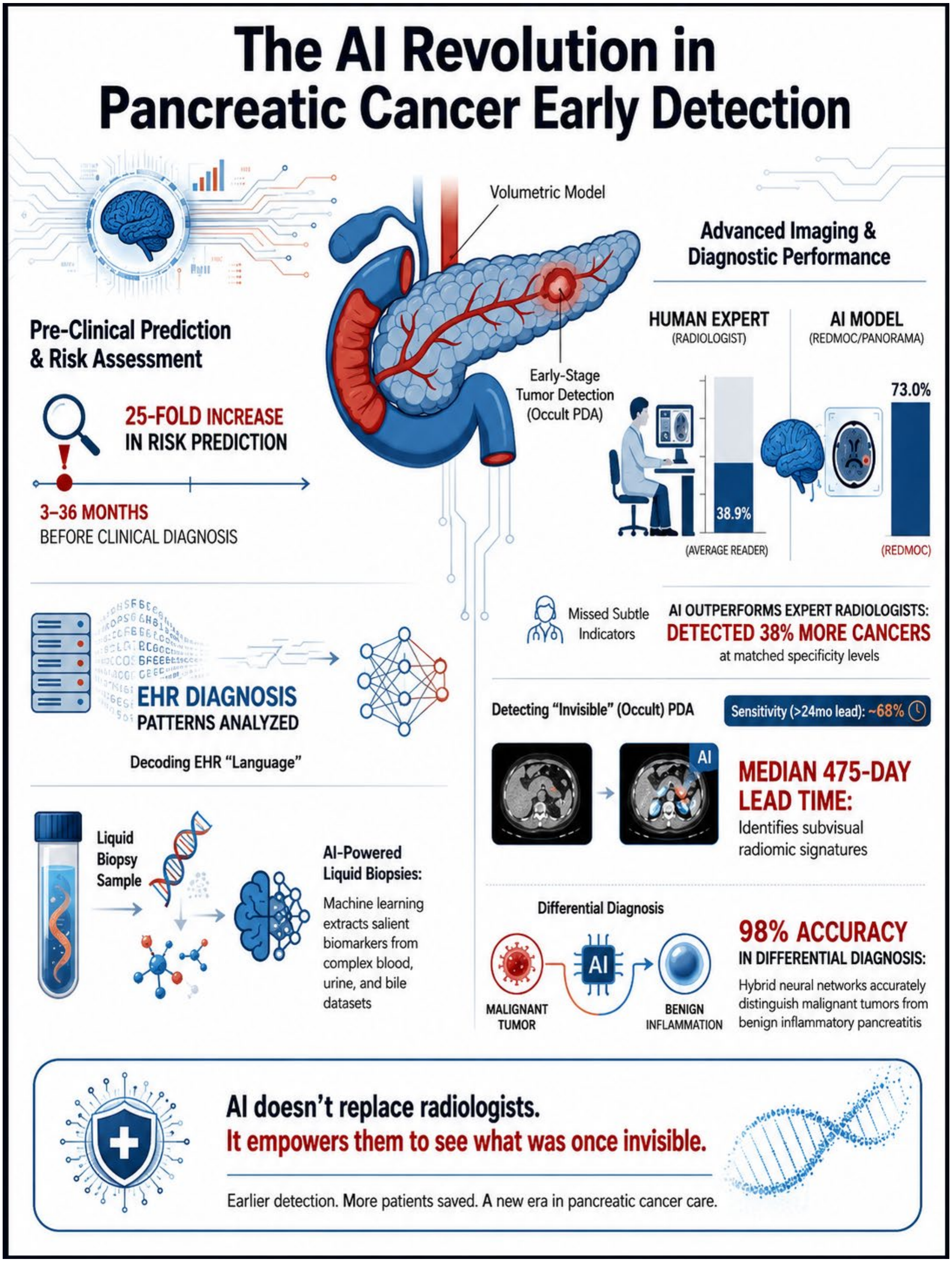
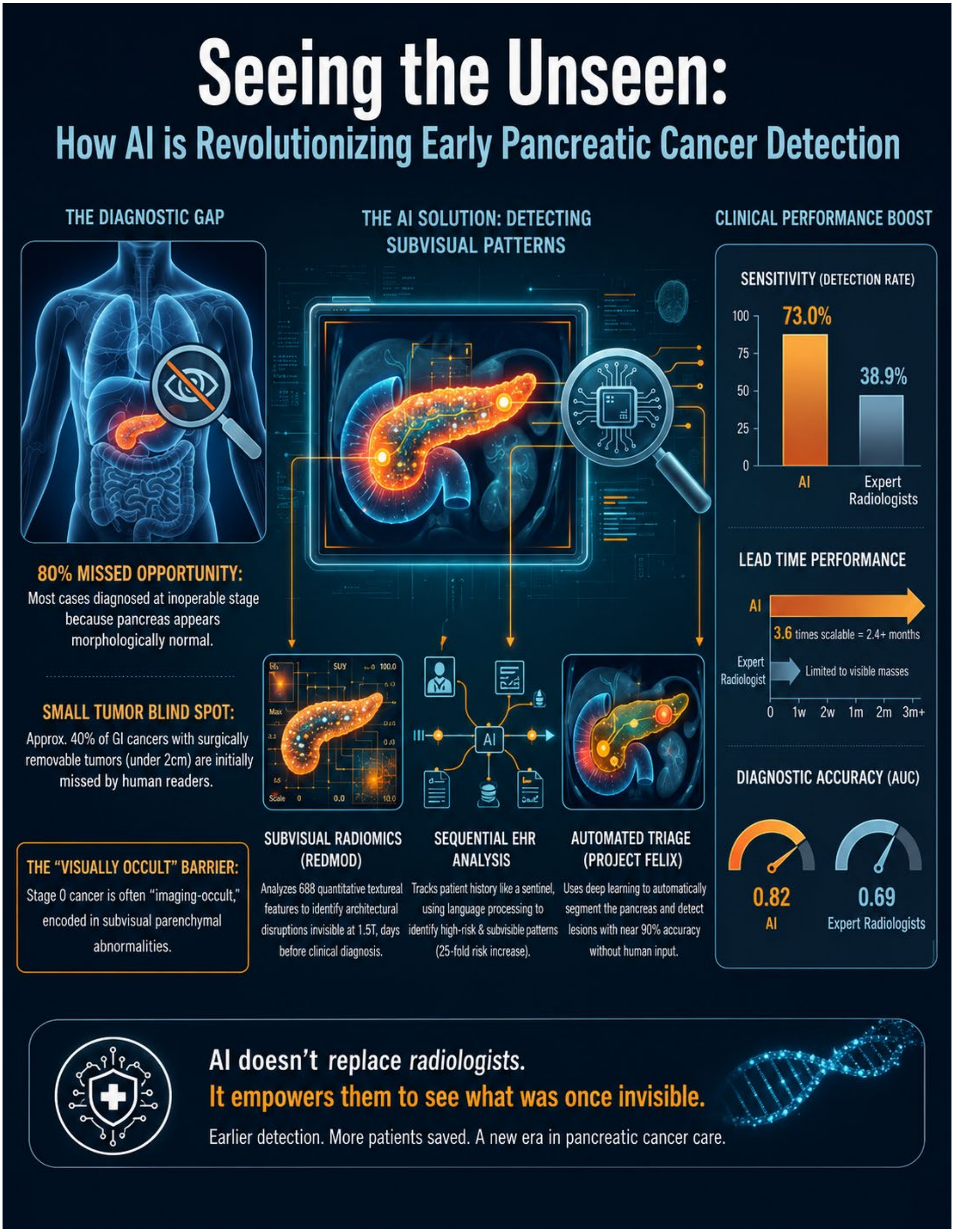


Figure 2 How AI early detection works



References

Kenner, B., Chari, S. T., Kelsen, D., Klimstra, D. S., Pandol, S. J., Rosenthal, M., Rustgi, A. K., Taylor, J. A., Yala, A., Abul-Husn, N., Andersen, D. K., Bernstein, D., Brunak, S., Canto, M. I., Eldar, Y. C., Fishman, E. K., Fleshman, J., Go, V. L. W., Holt, J. M., ... Wolpin, B. (2021). Artificial intelligence and early detection of pancreatic cancer: 2020 summative review. Pancreas, 50(3), 251-279. https://doi.org/10.1097/MPA.0000000000001762

Ku, J., Singhal, M., Burnette, M., & Hegazy, S. A. (2026). Machine learning and artificial intelligence in liquid biopsy-based early detection of pancreatic cancer: a scoping review. BJC Reports, 4(1), 26. https://doi.org/10.1038/s44276-026-00232-y

Mukherjee, S., Antony, A., Patnam, N. G., Trivedi, K. H., Karbhari, A., Bhinder, K. K., Zarrintan, A., Fletcher, J. G., Truty, M., Johnson, M. P., Chari, S. T., & Goenka, A. H. (2026). Next-generation AI for visually occult pancreatic cancer detection in a low-prevalence setting with longitudinal stability and multi-institutional generalisability. Gut, Online First. https://doi.org/10.1136/gutjnl-2025-337266

Prunoiu, V. M., Juverdeanu, O., Cosma, C., Laurentiu, S., Strâmbu, V., Petru, A. R., Stana, M., & Brătucu, M. N. (2026). Legal and ethical reflections on the use of artificial intelligence in the diagnosis and treatment of cancer: Who assumes responsibility? Frontiers in Artificial Intelligence, 9, Article 1812408. https://doi.org/10.3389/frai.2026.1812408

Udriştoiu, A. L., Cazacu, M. I., Gruionu, L. G., Gruionu, G., Iacob, A. V., Burtea, D. E., Ungureanu, B. S., Costache, M. I., Constantin, A., Popescu, C. F., Udriştoiu, Ş., & Săftoiu, A. (2021). Real-time computer-aided diagnosis of focal pancreatic masses from endoscopic ultrasound imaging based on a hybrid convolutional and long short-term memory neural network model. PLoS ONE, 16(6), Article e0251701. https://doi.org/10.1371/journal.pone.0251701

Zhang, L., Li, D., Su, T., Xiao, T., & Zhao, S. (2025). Effectiveness of radiomics-based machine learning models in differentiating pancreatitis and pancreatic ductal adenocarcinoma: Systematic review and meta-analysis. Journal of Medical Internet Research, 27(1), Article e72420. https://doi.org/10.2196/72420